

APPENDIX II: Yablo's Account of Mental Causation

In the spring of 1999 I was attempting to adapt Stephen Yablo's determination account of mental causation either to a setting in which determination relations are relativized to the physical laws or to a setting in which we posit "megawide events" (i.e., events that incorporate the physical laws as constituents). This attempt, if successful, would have yielded an account that is compatible with both strong and weak (i.e., nomic) supervenience. In the course of working on this, I happened upon a family of examples that seem to undermine the project. I then came to suspect that the threat to the project was a special case of a more general problem and that the same examples might threaten Yablo's original account of mental causation itself. The problem goes as follows

Virtually no one believes that a mere neurological event (e.g., the firing of certain neurons) taken in isolation from the larger brain to which it belongs is metaphysically sufficient for, say, x to think that A . The neurological event must be situated in a brain, and not just any brain but one that is operating in accordance with relevant physical laws (both synchronic and diachronic). Absent such laws, the brain would be "dead," and the brain-situated neurons would not be metaphysically sufficient for the occurrence of the mental event. It seems that the determination account can deal with this matter of physical laws in either of two ways.

The first is to posit, not just wide events (which include background conditions, such as temperature and pressure), but also "megawide events": events so wide that they include the laws of physics as constituents. If this alternative is adopted, however, the analogy to the drinking/guzzling example is lost (or at least weakened), for the latter determination relation holds without the inclusion of physical laws in the event of Socrates's guzzling. Furthermore, the "megawide event" approach does not seem very plausible: are there really *events* like this? (One could instead invoke megawide events that include as constituents dispositional properties that somehow code up physical laws (e.g., the property of being a body such that $f = ma$). This alternative seems to lead to the same outcome as the original.) But let us waive these two concerns, for this first approach seems to be beset by the more serious problem, which I am about to discuss.

The second approach the physical-laws problem is to *relativize* determination relations to the physical laws (or to associated dispositions that encode those laws). For example, (where L_π are the laws of physics): c is a determination of d relative to L_π iff, necessarily, given L_π , c occurs only if d occurs, but not conversely. With this relativized notion of determination in place, it seems that the account can proceed pretty much as it did before (and without having to appeal to megawide events).

This relativization of determination relations, however, appears to create a larger problem: it becomes too easy to generate "determination relations" between lawfully connected events. (This problem, *mutatis mutandis*, also it does so without having to posit megawide events.)

Consider a case of ordinary causal overdetermination in which c_1 causes e and c_2 also causes e . Suppose, in this connection, that the physical laws entail that if c_1 occurs so does e and that if c_2 occurs so does e . Now consider an otherwise similar case in which the physical laws are strengthened (call the strengthened laws L_{π^*}) so that they entail, in addition, that if c_1 occurs so does c_2 (but L_{π^*} does not entail the converse). Then, on the present relativized-determination proposal, it would follow that c_1 determines c_2 relative to L_{π^*} . For it is necessary that, given L_{π^*} , if c_1 occurs so does c_2 (but not conversely). Given this, the relativized-determination account directs us to apply the screening off test in order to decide whether c_1 or c_2 (or both) causes e . But in the nearest world(s) in which c_2 occurs in absence of c_1 , e would still occur, for in such worlds, c_2 would intuitively still be causally sufficient for e . (At the same time, running the test in the other

direction— c_1 without c_2 —is ruled out since c_1 determines c_2 relative to L_{π^*} .) Thus, we would be committed to holding that c_2 prevails over c_1 as the cause of e . But this is plainly mistaken: simply modifying the original case of ordinary overdetermination so that the physical laws also entail that c_1 is causally sufficient for c_2 would not undermine c_1 's status as a cause of e . If anything, c_1 ought to be identified as the cause of e in the envisaged situation. (This problem, if it is a problem, would not be not an isolated difficulty for the relativized solution; there would be a large class of problems having a similar underlying structure.)

Thus, we seem to be faced with a dilemma. To get the requisite determination relations in the case of mental causation, one must evidently relativize determination relations to the physical laws (or posit megawide events that include physical laws as constituents). But when this is done, too many "determination" relations are created, and the associated "determinables" are then wrongly identified as the cause of their target effects and the causal role of the "determinate" is wrongly dismissed. The reason is that, whenever a "determination" relation is present, one is obliged to apply only one half of the natural screening-off test—namely, to establish what happens when the "determinable" c_2 occurs in absence of the "determinate" c_1 . Of course, the target effect e will typically occur in such situations. However, if one were able to consider the other half of the test—for example, to establish what happens when c_1 occurs in absence of c_2 —we would see that the effect also occurs, thereby ruling out the unwanted conclusion that the "determinable" is *the* cause of e . In short, for such "determinate"/"determinable" cases, one half of the natural test is ruled out on the cheap, thereby creating the appearance that the "determinable" is the cause when it plainly is not. This leads to a worry: maybe this is all that is going on in the determination account of mental causation; maybe one half of the screening-off test is likewise being ruled out, creating the mere appearance of the mental event's being the cause. Now there are various ways to try to avoid this sort of worry within the determination framework, broadly construed; but as far as I can tell they only open the way for new instances of the same general sort of problem. (It also seems to me that this sort of difficulty would arise in the context of Shoemaker's account of mental causation in "Realization and Mental Causation" in terms of "causal-power inclusion"—unless one also adopts his strong necessitarian view of laws)

Perhaps a general moral can be drawn from the above problem. On the one hand, any test for causation, and in particular mental causation, must be run without somehow "building in" the physical laws (either by relativizing things to physical laws or by incorporating physical laws into megawide events), for otherwise one gets untoward consequences. But, on the other hand, in order to be even remotely plausible, metaphysical supervenience theses must "build in" laws in some such way. Taken together, these two observations suggest that an adequate general account of mental causation cannot rely on metaphysical supervenience in any essential manner. If so, accounting for mental causation is independent of the debate over supervenience, and so mental causation is unable to provide the missing theoretical justification of supervenience.